TITLE: Correlated evolution of androgen receptor and aromatase revisited

Submission: Letter

AUTHORS: Adam M. Reitzel and Ann M. Tarrant*

Biology Department
Woods Hole Oceanographic Institution
45 Water St.
Woods Hole, MA 02543

* Corresponding Author

email: atarrant@whoi.edu
phone: 508-289-3398
fax: 508-457-2134

KEYWORDS

Androgen Receptor, Aromatase, Correlated Evolution, Cytochrome p450, Nuclear Receptor

RUNNING TITLE

Correlated evolution of androgen receptor and aromatase revisited

Title length: 65

Abstract length: 239

Total length of text: 14, 039

Total page requirements: 1.7 pages text, 1 page figures plus legends, 0.3 page abstract, 0.5 page references

Number of references: 32

ABSTRACT

Conserved interactions among proteins or other molecules can provide strong evidence for coevolution across their evolutionary history. Diverse phylogenetic methods have been applied to identify potential coevolutionary relationships. In most cases, these methods minimally require comparisons of orthologous sequences and appropriate controls to separate effects of selection from the overall evolutionary relationships. In vertebrates, androgen receptor (AR) and cytochrome p450 aromatase (CYP19) share an affinity for androgenic steroids, which serve as receptor ligands and enzyme substrates. In a recent study, Tiwary and Li (2009) reported that AR and CYP19 displayed a signature of ancient and conserved interactions throughout all of the Eumetazoa (i.e., cnidarians, protostomes, and deuterostomes). Because these findings conflicted with a number of previous studies, we reanalyzed the data set used by Tiwary and Li. First, our analyses demonstrate that the invertebrate genes used in the previous analysis are not orthologous sequences, but instead represent a diverse set of nuclear receptors and cytochrome p450 enzymes with no confirmed or hypothesized relationships with androgens. Second, we show that (1) their analytical approach, which measures correlations in evolutionary distances between proteins, potentially led to spurious significant relationships due simply to conserved domains and (2) control comparisons provide positive evidence for a strong influence of evolutionary history. We discuss how corrections to this method and analysis of key taxa (e.g., duplications in the teleost fish and suiform lineages) can inform investigations of the coevolutionary relationships between androgen receptor and aromatase.

Molecular coevolution is the correlated evolution of two or more interacting molecules due to selection imposed by changes in each on the other. Molecular coevolution has been demonstrated in several cases where proteins directly interact, particularly when they form obligate complexes within molecular networks. These interactions have largely been elucidated by testing for correlated changes in amino acid or nucleic acid sequences using phylogenetic methods and/or structural models (Yeang and Haussler 2007; Pazos and Valencia 2008). Molecular coevolution of pairs of proteins may be more indirectly mediated through their interactions with a conserved third molecule, such as a co-factor, ligand, or substrate (McPartland, Norris, and Kilpatrick 2007). Particularly good candidates to investigate this later form of correlated evolution are components of steroid-signaling pathways, which require the actions of enzymes and receptors with specific, high-affinity interactions with hormones.

In a recent study, Tiwary and Li (2009) tested for correlated evolution of androgen receptor (AR) and aromatase (CYP19) throughout the animal lineage. AR is a ligand-activated member of the nuclear receptor superfamily (NR3C4) that specifically binds androgens. Aromatase, a cytochrome p450 enzyme (CYP19), catalyzes the synthesis of estrogens from androgen precursors. Thus, these two proteins share specificity for androgens and this indirect interaction may link their evolutionary and functional histories.

Tiwary and Li (2009) claimed to identify apparent AR and CYP19 genes with similarity searches throughout the Eumetazoa, including insects, a cnidarian, and other invertebrates. These authors reported a strong and significant correlation of protein distances between AR and CYP19, but not among background control proteins,

suggesting that AR and CYP19 have evolved at similar rates throughout most of animal evolution. Tiwary and Li (2009) concluded that AR and CYP are evolving in a correlated fashion, which they termed parallel evolution. However, to test hypotheses for correlated evolution, whether it be co- or parallel evolution, practitioners must use orthologous sequences. Confirmation of parallel evolution additionally requires identification of identical yet independent replacements in particular amino acids, which involve site-specific analyses in a phylogenetic context (Rokas and Carroll 2008). We show that their analysis violates the requirement for orthologous sequences, and further that the relationship is not strongly different from other surveyed proteins without a functional interaction.

The reported identification of AR and CYP19 across much of the animal kingdom by Tiwary and Li (2009) conflicts with previous literature. Phylogenetic studies of the nuclear receptor superfamily have shown that AR differentiated from an ancestral steroid receptor early in the vertebrate lineage (Thornton 2003; Bertrand et al. 2004; Bridgham et al. 2008). Similarly, CYP19 most likely evolved in the lineage leading to the cephalochordate-vertebrate ancestor (Campbell, Satoh, and Degnan 2004; Baker 2007). In a recent study, Markov et al. (2009) stated that in contrast to the report by Tiwary and Li, they found no evidence for an aromatase gene outside of the chordate lineage, but they did not explain the incongruent results.

We tested the evolutionary relationships of the nuclear receptors (NRs) and cytochrome p450s (CYPs) used by Tiwary and Li (2009) (see Supplemental File for all methods). Because their data set lacked designated outgroup sequences, we retrieved a set of sequences to represent the diversity within these superfamilies. We found that none of

the invertebrate sequences included in this earlier study are orthologous to either AR (fig. 1A) or CYP19 (fig. 1B). Our analysis strongly supports placement of the invertebrate NRs within diverse NR families, mostly in Nuclear Receptor Family 2. As previously reported (Holland et al. 2008; Schubert et al. 2008), the *Branchiostoma* steroid receptor was positioned as ancestral to the NR3 steroidogenic receptors. Similarly, the invertebrate CYPs used by Tiwary and Li (2009) represent diverse CYP families. As expected, the *Branchiostoma* CYP19 gene formed a strongly supported clade with the vertebrate CYP19s. All the vertebrate sequences were strongly supported as orthologs of AR and CYP19.

Analysis of the evolutionary relationships between AR and CYP19 is complicated because both of these proteins have been retained as duplicates in most teleosts (e.g., Chiang et al. 2001b; Ogino et al. 2009). The inclusion of a mixture of paralogs explains the unusual tree topology for fish ARs discussed by Tiwary and Li (2009). By selecting the least divergent paralogs, we recovered a topology reflecting the evolutionary relationships of these fish (not shown).

In spite of using a mix of orthologs and distantly related homologs (i.e., members of the same superfamily), Tiwary and Li (2009) obtained significant regressions when comparing pairwise molecular distances of AR and CYP19 among the sampled taxa. A correlation of protein distances between pairs of orthologous sequences from two or more species is expected even in the absence of coevolution because the evolutionary time between species is identical for all genes. In an effort to correct for shared evolutionary history, the authors performed similar analyses for a number of “background random proteins” from which they reported no significant relationships, suggesting that shared

evolutionary history could not account for the strong correlations between AR and CYP19. We analyzed these relationships in two ways: 1) comparing the molecular distance data separately for orthologous and homologous sequences, and 2) evaluating the effectiveness of background control comparisons in interpreting these regressions.

We conducted three sets of regression analyses to test for correlations between protein distances for AR and CYP19: 1) all sequences, 2) orthologous vertebrate sequences, and 3) the invertebrate sequences (none of which are orthologs of either AR or CYP19). For the full data set, we obtained a strong and significant correlation between AR and CYP19 identical to the one reported by Tiwary and Li (fig. 2A). When we compared only orthologs, we obtained a slightly poorer fit with a highly significant correlation (fig. 2B). Finally, when we regressed the invertebrate sequences, we obtained a weak but still significant regression (fig. 2C).

To more fully test the role of background control genes in interpreting the correlation of molecular distances between AR and CYP19, we performed additional comparisons using a reduced set of orthologs from eight vertebrate species distributed throughout the subphylum. By reducing the number of species, we may lose power in the ability to detect significant relationships. Thus, there is the possibility of committing a Type 2 error, but any significant relationships we detect are robust and would presumably only strengthen with the addition of more taxa. When we compare the regression of AR and CYP19 distances in the reduced data set, we obtained a significant relationship (fig. 3A), which is a better fit than the broader data set, due principally to the removal of the divergent fish paralogs.

We tested two sets of control proteins for background evolutionary relationships. First, we selected two genes from the same superfamilies, the germ cell nuclear factor (GCNF, an NR) and CYP51. We selected these genes because they are not present as duplicates within the vertebrates surveyed, and because they have no reported interactions with one another or with AR or CYP19.

When we regressed protein distances for GCNF and CYP51, we obtained a significant correlation between these two presumably non-interacting proteins (fig. 3B). We also obtained significant regressions when we compared AR and CYP51 ($r^2 = 0.57$, $p < 0.0001$) and GCNF and CYP19 ($r^2 = 0.77$, $p < 0.0001$) (supp. fig. 3). Thus, if we used only significant correlations when inferring coevolution of proteins, we would conclude that there is a significant functional relationship between any of the NR – CYP pairs we tested. There is no evidence, experimental or otherwise, to hypothesize this is the case.

In the second comparison we used four (glucagon, myoglobin, erythropoietin, and glucokinase) of the background proteins selected by Tiwary and Li. They originally compared six background proteins among one another, but not with either AR or CYP19. Two of these control proteins (amylase, ferritin) were inappropriate for our analysis, and we argue are poor choices for background comparisons, because they have undergone lineage-specific duplications and divergence in various animals (e.g., amylase, Meisler and Ting 1993; ferritin, Colbourne et al. 2007). For the other four genes, we observed strong and significant correlations between each control gene and AR (erythropoietin: $r^2 = 0.96$, $p < 0.0001$, fig. 3C; glucagon: $r^2 = 0.83$, $p < 0.0001$; myoglobin: $r^2 = 0.75$, $p = 0.0001$; glucokinase: $r^2 = 0.50$, $p = 0.0016$). We observed strong correlations for erythropoietin and myoglobin even though we could not include sequences from some

species due to independent gene loss (e.g., myoglobin lost in *Xenopus* and other anurans (Fuchs, Burmester, and Hankeln 2006; Xi et al. 2007)) or absence of a gene model (erythropoietin from *Xenopus* or *Gallus*). Thus, the poor correlations for comparisons using these genes reported by Tiwary and Li likely represent their mixing of orthologs with nonorthologous sequences, similar to their analyses with AR and CYP19. When we regressed molecular distances for glucokinase, the correlation was considerably strengthened (and similar to the relationship between AR and CYP19), when we removed *Xenopus* ($r^2 = 0.92$, $p < 0.0001$), which has a divergent glucokinase sequence that resulted in relatively long branch in the phylogenetic tree (supp. fig. 4). Although the fit for the relationships with the control genes are similar to or weaker than for the original AR – CYP19 comparison, the significant correlations in all comparisons suggest that the previous controls by Tiwary and Li were not adequate control comparisons and instead reflect a significant signal of the evolutionary relationships between species. These data suggest that analysis of the correlated evolution between AR and CYP19 should include a correction for the background phylogenetic relationships using a suite of non-interacting proteins from diverse gene families (e.g., apply a 'phylogenetic vector' as a correction factor", Sato et al. 2005).

An alternative and more direct approach to study co-evolution of AR and CYP19 could combine experimental and molecular modeling approaches to study compensatory mutations for conserved ligand binding throughout vertebrates. Crystal structures have been determined for human AR (Pereira de Jésus-Tran et al. 2006) and CYP19 (Ghosh et al. 2009). Critical residues for binding of androgens have been identified for each protein (AR: (Matias et al. 2000; Pereira de Jésus-Tran et al. 2006); CYP19: (Ghosh et al. 2009)).

From a preliminary analysis of AR, only one of these positions (residue 749 in human AR) varied in more than one species, in this case between fish and tetrapods. An analysis of the AR crystal structure shows only weak interactions of this residue with either testosterone or DHT (Pereira de Jésus-Tran et al. 2006); it seems unlikely that this amino acid difference exerts a strong functional change. Similarly, for CYP19, only amino acid positions 372 and 373, both in the catalytic cleft, show differences in more than one taxon, with these positions conserved among tetrapods but not with fishes. Thus, there are few replacements among the potentially critical residues for each of these proteins. While the critical contact residues are highly conserved, the relative affinity of steroidal androgens for AR varies among taxa. The primary physiological ligand(s) are dihydrotestosterone in mammals, 11-ketotestosterone and testosterone in fishes, and androstenedione in the sea lamprey (Sperry and Thomas 1999a; Sperry and Thomas 2000; Bryan, Scott, and Li 2007). Together, these results indicate that integrative studies of AR protein structure and function will be necessary for understanding the evolution of ligand binding by AR.

Additionally, a comparative approach could take advantage of independent duplications of CYP19 in the Suiformes (Corbin et al. 2007; Conley, Corbin, and Hughes 2009) and AR and CYP19 in teleosts (Chiang et al. 2001a; Douard et al. 2008). In the case of AR, biochemical studies have characterized two distinct proteins (termed AR1 and AR2) in teleosts that differ in expression and ligand affinity (Sperry and Thomas 1999b; Sperry and Thomas 2000). Similarly, CYP19 paralogs have distinct expression patterns in teleosts (Callard et al. 2001; Chiang et al. 2001b), and expression patterns and substrate affinity in the Suiform lineage (Corbin et al. 2007; Conley, Corbin, and Hughes

2009). Further functional characterization of duplicated AR and CYP19 genes will help to elucidate how these proteins have diversified and whether evolutionarily related paralogs have retained conserved functions.

In conclusion, the original analysis by Tiwary and Li (2009) that detected correlated evolution of AR and CYP19 throughout the Eumetazoa included a mixture of orthologs and more distantly related homologs and thus a critical criterion for assessing coevolutionary change in proteins was violated. Additional regression analyses show that a strong signal of evolutionary history and overall conservation of particular domains likely led to spurious significant relationships. Future studies of potential correlated evolution of AR and CYP19 should include molecular distance corrections that remove the species-level evolutionary relationships, functional studies of duplications, and comparative protein modeling.

FUNDING

AMR was supported by the Postdoctoral Scholar Program at the Woods Hole Oceanographic Institution, with funding provided by The Beacon Institute for Rivers and Estuaries, and AMT was supported by WHOI Assistant Scientist Endowed Support.

ACKNOWLEDGEMENTS

We thank Mark Hahn and his lab members for productive discussions. We would also like to thank the handling editor and four referees for thoughtful reviews that significantly improved our study.

FIGURE LEGENDS

Figure 1. Maximum-likelihood analysis of the NRs (A) and CYPs (B) used by Tiwary and Li (2009) with additional outgroup sequences. Full trees showing all terminal branches are presented in supp. fig.1 and 2. (A) Vertebrate ARs form a strongly supported clade (bootstrap = 95) and are positioned with other steroid-binding receptors in NR family 3. However, the invertebrate sequences used by Tiwary and Li (2009) are distributed throughout the NR superfamily (underlined). (B) Vertebrate aromatase (CYP19) sequences form a clade with strong support (bootstrap = 100). However, the invertebrate sequences used by Tiwary and Li (2009) are distributed throughout the CYP superfamily (underlined).

Figure 2. Regressions of pairwise distances for NRs and CYPs (A) Protein distances for all sequences used by Tiwary and Li (2009) are strongly and significantly correlated, with values identical to those previously reported. (B) Distances for AR and CYP19 orthologs (i.e., vertebrate sequences) were also strongly and significantly correlated, with a slightly poorer fit. (C) Distances for the invertebrate NRs and CYPs, none of which are closely related to AR or CYP19, were also positively and significantly correlated, although the fit is weaker.

Figure 3. Regressions of pairwise protein distances between (A) AR and CYP19, (B) GCNF and CYP51, and (C) AR and erythropoietin from eight vertebrates. All protein pairs were significantly correlated, suggesting that the correlation of molecular distances is largely due to evolutionary history, not positive selection (additional comparisons in supp. fig. 3).

REFERENCES

Baker, M. E. 2007. Amphioxus, a primitive chordate, is on steroids: evidence for sex steroids and steroidogenic enzymes. Endocrinology **148**:3551-3553.

Bertrand, S., F. G. Brunet, H. Escriva, G. Parmentier, V. Laudet, and M. Robinson-Rechavi. 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. Molecular Biology and Evolution **21**:1923-1937.

Bridgham, J. T., J. E. Brown, A. Rodriguez-Mari, J. M. Catchen, and J. W. Thornton. 2008. Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. PLoS Genetics **4**:e1000191.

Bryan, M. B., A. P. Scott, and W. Li. 2007. The sea lamprey (*Petromyzon marinus*) has a receptor for androstenedione. Biology of Reproduction **77**:688-696.

Callard, G. V., A. V. Tchoudakova, M. Kishida, and E. Wood. 2001. Differential tissue distribution, developmental programming, estrogen regulation and promoter characteristics of cyp19 genes in teleost fish. The Journal of Steroid Biochemistry and Molecular Biology **79**:305-314.

Campbell, R. K., N. Satoh, and B. M. Degnan. 2004. Piecing together evolution of the vertebrate endocrine system. Trends in Genetics **20**:359-366.

Chiang, E. F.-L., Y.-L. Yan, Y. Guiguen, J. Postlethwait, and B. Chung. 2001a. Two cyp19 (p450 aromatase) genes on duplicated zebrafish chromosomes are expressed in ovary or brain. Molecular Biology and Evolution **18**:542-550.

Chiang, E. F.-L., Y.-L. Yan, S.-K. Tong, P.-H. Hsiao, Y. Guiguen, J. Postlethwait, and B.-C. Chung. 2001b. Characterization of duplicated zebrafish cyp19 genes. Journal of Experimental Zoology **290**:709-714.

Colbourne, J., B. Eads, J. Shaw, E. Bohuski, D. Bauer, and J. Andrews. 2007. Sampling Daphnia's expressed genes: preservation, expansion and invention of crustacean genes with reference to insect genomes. BMC Genomics **8**:217.

Conley, A. J., C. J. Corbin, and A. L. Hughes. 2009. Adaptive evolution of mammalian aromatases: lessons from Suiformes. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology **311A**:346-357.

Corbin, C., A. Hughes, J. Heffelfinger, T. Berger, T. Waltzek, J. Roser, T. Santos, M. Miglino, M. Oliveira, F. Braga, F. Meirelles, and A. Conley. 2007. Evolution of suiform aromatases: ancestral duplication with conservation of tissue-specific expression in the collared peccary (*Pecari tayassu*). Journal of Molecular Evolution **65**:403-412.

Douard, V., F. Brunet, B. Boussau, I. Ahrens-Fath, V. Vlaeminck-Guillem, B. Haendler, V. Laudet, and Y. Guiguen. 2008. The fate of the duplicated androgen receptor in fishes: a late neofunctionalization event? BMC Evolutionary Biology **8**:336.

Fuchs, C., T. Burmester, and T. Hankeln. 2006. The amphibian globin gene repertoire as revealed by the Xenopus genome. Cytogenetic and Genome Research **112**:293-306.

Ghosh, D., J. Griswold, M. Erman, and W. Pangborn. 2009. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. Nature **457**:219-223.

Holland, L. Z., R. Albalat, K. Azumi, E. Benito-Gutierrez, M. J. Blow, M. Bronner-Fraser, F. Brunet, T. Butts, S. Candiani, L. J. Dishaw, D. E. K. Ferrier, J. Garcia-

Fernandez, J. J. Gibson-Brown, C. Gissi, A. Godzik, F. Hallbook, D. Hirose, K. Hosomichi, T. Ikuta, H. Inoko, M. Kasahara, J. Kasamatsu, T. Kawashima, A. Kimura, M. Kobayashi, Z. Kozmik, K. Kubokawa, V. Laudet, G. W. Litman, A. C. McHardy, D. Meulemans, M. Nonaka, R. P. Olinski, Z. Pancer, L. A. Pennacchio, M. Pestarino, J. P. Rast, I. Rigoutsos, M. Robinson-Rechavi, G. Roch, H. Saiga, Y. Sasakura, M. Satake, Y. Satou, M. Schubert, N. Sherwood, T. Shiina, N. Takatori, J. Tello, P. Vopalensky, S. Wada, A. Xu, Y. Ye, K. Yoshida, F. Yoshizaki, J.-K. Yu, Q. Zhang, C. M. Zmasek, P. J. de Jong, K. Osoegawa, N. H. Putnam, D. S. Rokhsar, N. Satoh, and P. W. H. Holland. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Research **18**:1100-1111.

Markov, G. V., R. Tavares, C. Dauphin-Villemant, B. A. Demeneix, M. E. Baker, and V. Laudet. 2009. Independent elaboration of steroid hormone signaling pathways in metazoans. Proceedings of the National Academy of Sciences **106**:11913-11918.

Matias, P. M., P. Donner, R. Coelho, M. Thomaz, C. Peixoto, S. Macedo, N. Otto, S. Joschko, P. Scholz, A. Wegg, S. Basler, M. Schafer, U. Egner, and M. A. Carrondo. 2000. Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. Journal of Biological Chemistry **275**:26164-26171.

McPartland, J. M., R. W. Norris, and C. W. Kilpatrick. 2007. Coevolution between cannabinoid receptors and endocannabinoid ligands. Gene **397**:126-135.

Meisler, M. H., and C.-N. Ting. 1993. The remarkable evolutionary history of the human amylase genes. Critical Reviews in Oral Biology & Medicine **4**:503-509.

Ogino, Y., H. Katoh, S. Kuraku, and G. Yamada. 2009. Evolutionary history and functional characterization of androgen receptor genes in jawed vertebrates. Endocrinology:en.2009-0523.

Pazos, F., and A. Valencia. 2008. Protein co-evolution, co-adaptation and interactions. EMBO Journal **27**:2648-2655.

Pereira de Jésus-Tran, K., P.-L. Côté, L. Cantin, J. Blanchet, F. Labrie, and R. Breton. 2006. Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. Protein Science **15**:987-999.

Rokas, A., and S. B. Carroll. 2008. Frequent and widespread parallel evolution of protein sequences. Molecular Biology and Evolution **25**:1943-1953.

Sato, T., Y. Yamanishi, M. Kanehisa, and H. Toh. 2005. The inference of protein-protein interactions by co-evolutionary analysis is improved by excluding the information about the phylogenetic relationships. Bioinformatics **21**:3482-3489.

Schubert, M., F. Brunet, M. Paris, S. Bertrand, G. Benoit, and V. Laudet. 2008. Nuclear hormone receptor signaling in amphioxus. Development Genes and Evolution **218**:651-665.

Sperry, T. S., and P. Thomas. 1999a. Characterization of two nuclear androgen receptors in Atlantic croaker: comparison of their biochemical properties and binding specificities. Endocrinology **140**:1602-1611.

Sperry, T. S., and P. Thomas. 2000. Androgen binding profiles of two distinct nuclear androgen receptors in Atlantic croaker (*Micropogonias undulatus*). The Journal of Steroid Biochemistry and Molecular Biology **73**:93-103.

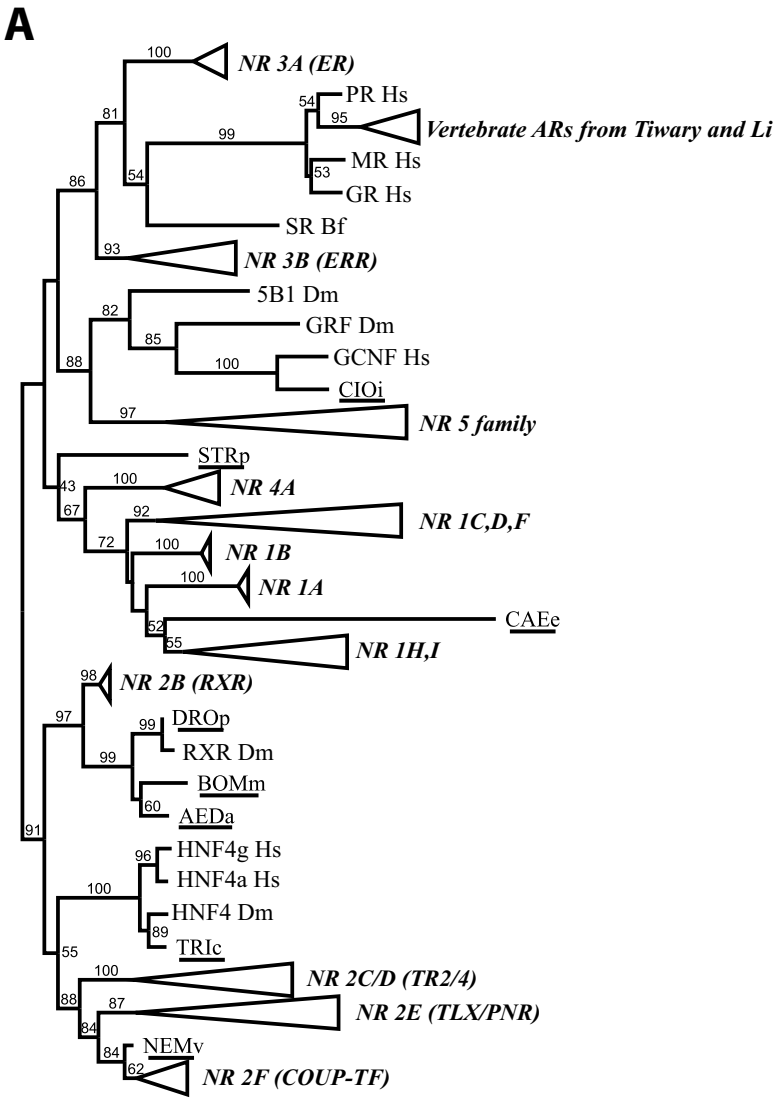
Sperry, T. S., and P. Thomas. 1999b. Identification of two nuclear androgen receptors in kelp bass (*Paralabrax clathratus*) and their binding affinities for xenobiotics: comparison with Atlantic croaker (*Micropogonias undulatus*) androgen receptors. Biology of Reproduction **61**:1152-1161.

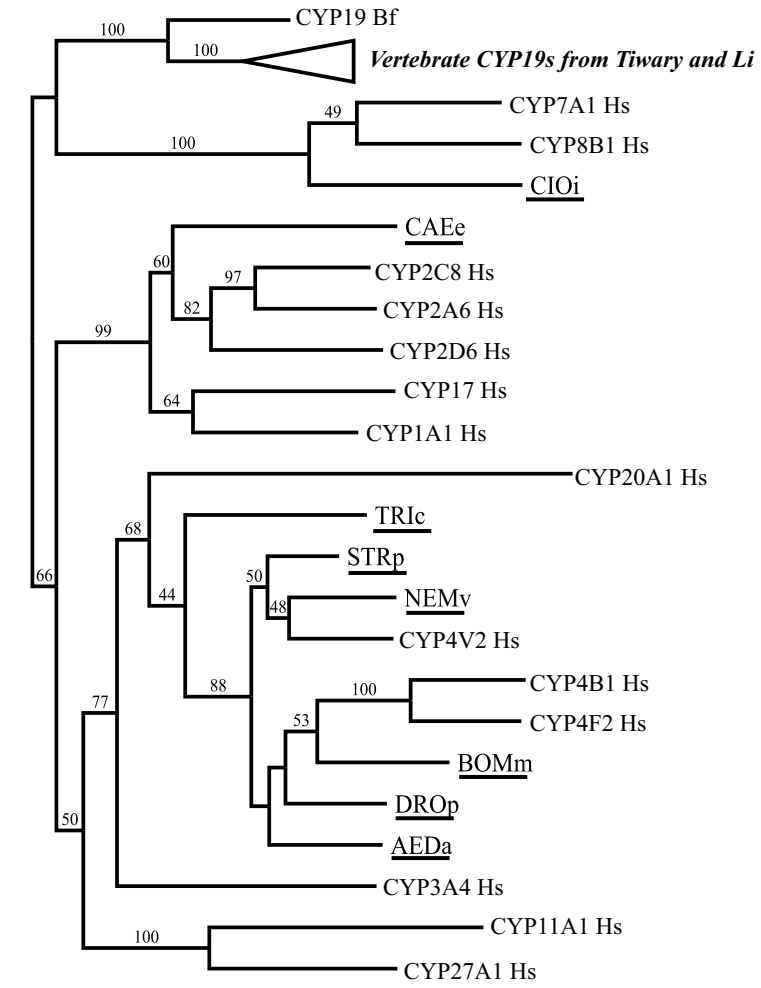
Thornton, J. W. 2003. Nuclear receptor diversity: phylogeny, evolution and endocrine disruption. Pure and Applied Chemistry **75**:1827-1839.

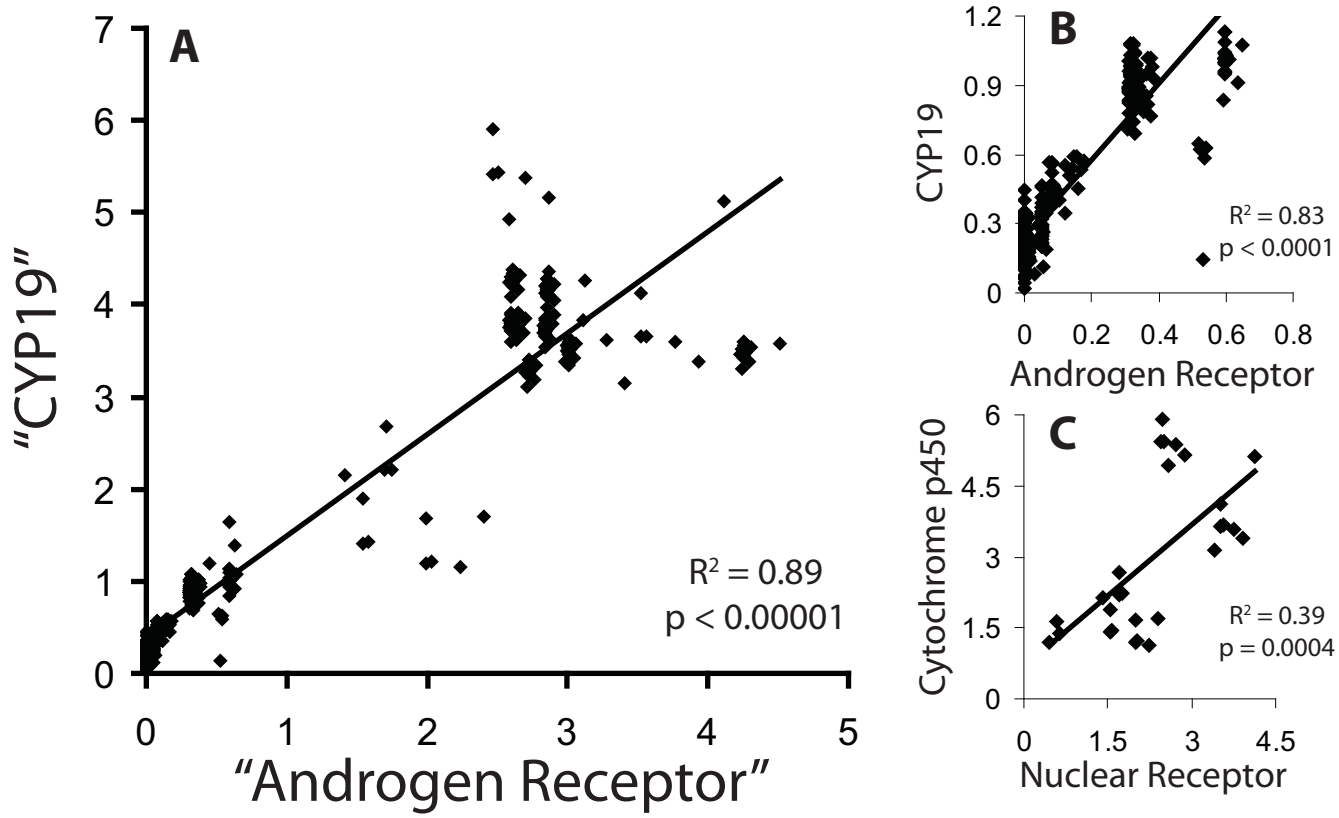
Tiwary, B. K., and W.-H. Li. 2009. Parallel evolution between aromatase and androgen receptor in the animal kingdom. Molecular Biology and Evolution **26**:123-129.

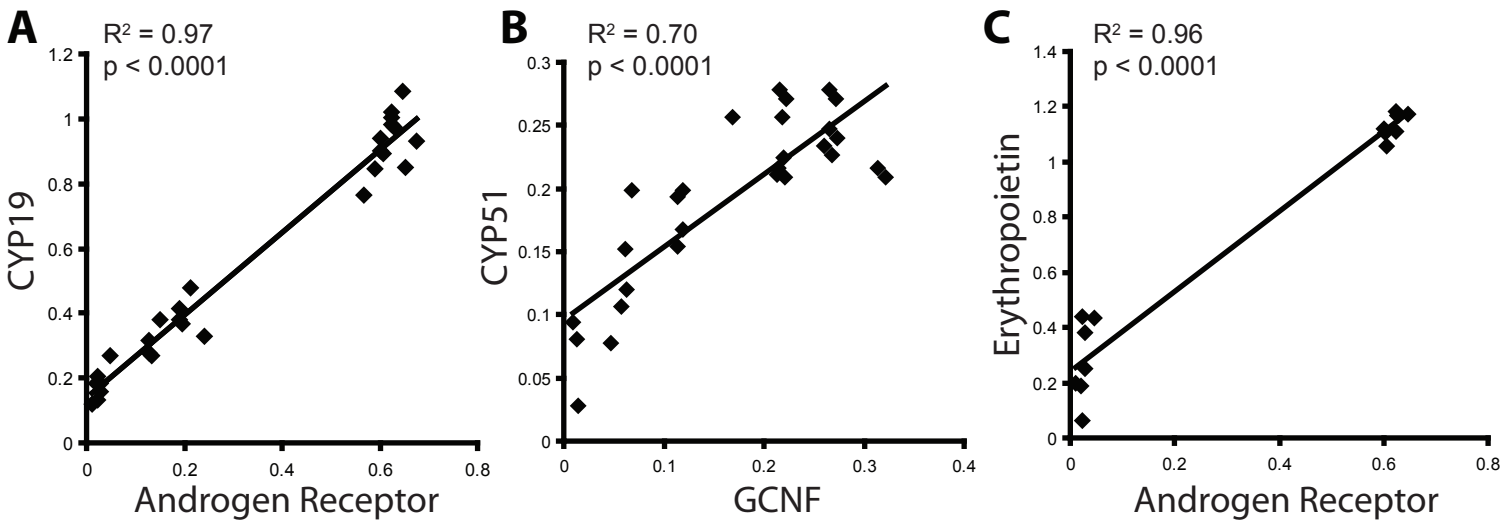
Xi, Y., M. Obara, Y. Ishida, S. Ikeda, and K. Yoshizato. 2007. Gene expression and tissue distribution of cytoglobin and myoglobin in the Amphibia and Reptilia: Possible compensation of myoglobin with cytoglobin in skeletal muscle cells of anurans that lack the myoglobin gene. Gene **398**:94-102.

Yeang, C.-H., and D. Haussler. 2007. Detecting coevolution in and among protein domains. PLoS Computational Biology **3**:e211.


A
NR 3A (ER)
PR Hs
Vertebrate ARs from Tiwary and Li
MR Hs
GR Hs
SR Bf
NR 3B (ERR)
5B1 Dm
GRF Dm
GCNF Hs
CIOi
NR 5 family
STRp
NR 4A
NR 1C,D,F
NR 1B
NR 1A
CAEe
NR 1H,I
NR 2B (RXR)
DROp
RXR Dm
BOMm
AEDa
HNF4g Hs
HNF4a Hs
HNF4 Dm
TRIc
NR 2C/D (TR2/4)
NR 2E (TLX/PNR)
NEMv
NR 2F (COUP-TF)


B
CYP19 Bf
Vertebrate CYP19s from Tiwary and Li
CYP7A1 Hs
CYP8B1 Hs
CIOi
CAEe
CYP2C8 Hs
CYP2A6 Hs
CYP2D6 Hs
CYP17 Hs
CYP1A1 Hs
CYP20A1 Hs
TRIc
STRp
NEMv
CYP4V2 Hs
CYP4B1 Hs
CYP4F2 Hs
BOMm
DROp
AEDa
CYP3A4 Hs
CYP11A1 Hs
CYP27A1 Hs


A
"CYP19"
"Androgen Receptor"
R² = 0.89
p < 0.00001
B
CYP19
Androgen Receptor
R² = 0.83
p < 0.0001
C
Cytochrome p450
Nuclear Receptor
R² = 0.39
p = 0.0004


A
R² = 0.97
p < 0.0001
CYP19
Androgen Receptor
B
R² = 0.70
p < 0.0001
CYP51
GCNF
C
R² = 0.96
p < 0.0001
Erythropoietin
Androgen Receptor